|  |
| --- |
| Apache Lucene Tutorial -1 | Solr Search |

Graphical user interface, application

Description automatically generated

Sorl is enterprise grade web search platform built on top of lucene

**Index document and make it searchable**

Lucene takes document and index the documents.

Crawlers can take documents from the web and give data to lucene . And lucene can index the received data.

Doug Cutting “Creator” of lucene.

A picture containing diagram

Description automatically generated

|  |  |
| --- | --- |
| Indexing and Searching |  |
| Ranked searching (best results returned first) | Wild cards or prefix queries or Boolean queries. |
| Many powerful query types: phrase queries, wildcard queries, proximity queries |  |
| Field searching (e.g. title, author, contents etc.) |  |
| Sorting by any field |  |
| Multiple index searching with merged results |  |
| Flexible faceting, highlighting, joins and results grouping |  |

Block indexing or incremental indexing—it accumulates all the index and updates them in incremental order it does not cause changes in immediate changes.

Diagram

Description automatically generated

|  |
| --- |
| The first step of all search engines, is a concept called indexing |
| Index: is the process of original data into a highly efficient cross-reference lookup order to facilitate rapid searching |
| Analyze: Search engine does not index text directly. The text are broken into a series of individual atomic elements called tokens. |
| Searching: is the process of consulting the search index and retrieving the documents matching the query, sorted in the request sort order. |
|  |